EFFECTS OF DIELECTRIC HETEROGENEITY IN THE PERFORMANCE OF BREAST TUMOUR CLASSIFIERS

R. C. Conceição, M. O’Halloran, M. Glavin, and E. Jones †

Electrical & Electronic Engineering
College of Engineering and Informatics
National University of Ireland Galway, Ireland

Abstract—Breast cancer detection using Ultra Wideband Radar has been thoroughly investigated over the last decade. This breast imaging modality is based on the dielectric properties of normal and cancerous breast tissue at microwave frequencies. However, the dielectric properties of benign and malignant tumours are very similar, so tumour classification based on dielectric properties alone is not feasible. Therefore, classification methods based on the Radar Target Signature of tumours need to be further developed to classify tumours as either benign or malignant. Several studies have addressed the issue of tumour classification based on the size, shape and surface texture of the tumour. In general, these studies examined the performance of classification algorithms in primarily dielectrically homogeneous breast models. These relatively simplistic models do not provide a realistic test platform for the evaluation of tumour classification algorithms. This paper examines the classification of tumours under realistic dielectrically heterogeneous conditions. Four different heterogeneous scenarios are considered, with varying levels of heterogeneity and complexity. In this paper, the performance and robustness of tumour classification algorithms under these realistic conditions are examined and discussed.

Received 24 December 2010, Accepted 15 February 2011, Scheduled 17 February 2011
Corresponding author: Raquel Cruz Conceicao (raquelcruzconceicao@gmail.com).
† All are also with Bioelectronics Research Cluster, NCBES, National University of Ireland Galway, Ireland.
1. INTRODUCTION

Diagnosis of breast cancer has been investigated using Microwave Imaging (MI) for over a decade [1–13]. MI is a promising alternative to X-Ray Mammography, currently the most commonly used breast imaging modality. In this context, MI is also an appealing imaging method since it does not require the compression of the breast and uses non-ionising radiation. The issue of accurate tumour classification is of considerable clinical importance in order to reduce the incidence of false-negative and false-positive diagnosis [6, 14, 15], reducing unnecessary costs on the health system and distress to patients. While recent MI techniques may be sufficient to detect the presence of a tumour, further classification algorithms have to be developed to obtain more accurate information on the tumour type.

In previous studies, such as Davis et al. [16], the breast tissues and the tumours are modelled using a 3D Total-Field/Scattered-Field (TF/SF) Finite Difference Time Domain (FDTD) model in which the breast tissue is considered primarily homogeneous. This methodology ensures that the tumour response is isolated in the backscattered signals response, allowing for the simple extraction of the Radar Target Signature of the tumour and its subsequent classification.

In previous studies by the authors [17–21], the homogeneous models developed by Davis et al. [16] were replicated and new feature extraction methods were investigated, such as Independent Component Analysis (ICA) and Discrete Wavelet Transform (DWT), while various classifiers were also compared, including Quadratic Discriminant Analysis (QDA), Support Vector Machines (SVM) and Spiking Neural-Networks (SNNs). These classifiers were applied to the selected features by following up to eight different multi-stage classification architectures.

In studies from Chen et al. [22–25] and Teo et al. [26], some heterogeneity is introduced into the breast phantom by modelling areas of clutter around the tumour models. The approach of these studies is to extract and process the late-time or early-time responses of the backscatter breast response so that tumour classification can be accomplished. In these studies, 2D FDTD simulations are completed and a database of tumour models is used, comprising of four different tumour shapes at a constant size.

In this paper, the effect of dielectric heterogeneity within the breast tissue is further examined by introducing a cluster of fibroglandular breast tissue extracted from 3D MRI models of the breast, taken from the UWCEM Numerical Breast Phantom Repository [27]. The analysis pursued in this paper follows a similar
structure to the authors’ previous papers [17,18,20,21] as a 3D
database of tumours of different sizes and shapes based on Gaussian
Random Spheres (GRSs) are again modelled with dielectric properties
as reported by Lazebnik et al. [28,29]. However, the inclusion of
fibroglandular tissue in the breast models in this paper presents a
much more challenging scenario for classification analysis than the
homogeneous breast models previously used.

The contributions of this paper are as follows:

• Use of a larger database with 480 tumour models comprising 30
  models for each of the four sizes combined with the four shapes.

• Evaluation of one feature extraction method, Principal Compo-
  nent Analysis (PCA), with a classification method, Support Vector
  Machines (SVM), as well as with the following classification ar-
  chitectures: Fine-Size-Coarse-Shape (FSCS) and Fine-Size-Fine-
  Shape (FSFS), as these combinations provided better classification
  performance in the previous authors’ studies [17, 18].

• Creation of four different dielectrically heterogeneous breast
  scenarios: (I) breast model with a cluster of fibroglandular breast
  tissue in a fixed location independent of the tumour location (as
  used in [20]); (II) breast model with a cluster of fibroglandular
  breast tissue, in a fixed location, possibly overlapping with
  tumour; (III) breast model in which one cluster of fibroglandular
  breast tissue is modelled across a range of random different
  locations within the breast; and finally, (IV) breast model in which
  two clusters of fibroglandular tissue are modelled across a range
  of random different locations within the breast.

2. NUMERICAL SIMULATION

The numerical simulation for this paper is similar to that described
in previous papers [17–21]. Therefore, a brief overview of the main
considerations is presented here.

The tumour models used in this study are based on GRSs [30,31],
similar to the method used in [17–21]. A database of 480 tumour
models was created with the purpose of replicating early-stage tumours
(less than 10 mm in radius). The shapes vary between smooth,
macrolobulated (benign tumours), microlobulated and spiculated
(malignant tumours) and the average sizes of the radii of the spheres
take discrete values: 2.5, 5, 7.5 and 10 mm.

A 3D Finite-Difference Time-Domain (FDTD) model is used
to model the dielectric properties of the breast and the tumour
tissues. The FDTD model has a 0.5 mm cubic grid resolution. The
backscattered signals are generated through a Total-Field/Scattered-Field (TF/SF) region terminated by a Uniaxial Perfectly Matched Layer (UPML) which suppresses any boundary reflections. The tumours and the fibroglandular clusters included in this study are both completely embedded in the Total Field (TF) [16–21].

All simulations are repeated for different heterogeneous scenarios, in which one or two fibroglandular clusters are introduced at different locations. The inclusion of clusters of fibroglandular breast tissue in the breast models allow for the analysis of the effect of dielectric heterogeneity on the tumour classification methods. This is an important consideration as the inclusion of heterogeneity permits a more realistic test platform for the classification algorithms. The simulations are completed as follows:

- The Total Field (TF) is located at the centre of the Scattered Field (SF) and is represented by a cube with 50 mm on each side. Figure 1 shows a representation of the TF/SF grid, with the location of the origin of the first incident plane wave and observer point (filled circle) as well as the position of the tumour. All four observation points are represented by small circles.

- A cluster of fibroglandular breast tissue is a block volume of 1 cm$^3$ representing a well-defined fibroglandular structure taken from a breast model from the UWCEM Numerical Breast Phantom Repository [27]. A number of clusters are located within the TF region at different locations, depending on the specific breast model.

The block of fibroglandular breast tissue is extracted from a geometrically and dielectrically accurate 3D breast model provided by the UWCEM Numerical Breast Phantom Repository at the University of Wisconsin repository [27], which provides the spatial distribution of the different constituent tissues within the breast. These are 3D MRI-derived models taken from patients lying in the prone position. The different tissues within the breast are mapped to the dielectric properties from Lazebnik et al. [28, 29]. For this study, a 1 cm$^3$ grid is extracted from a breast phantom available in [27] (phantom ID 071904). This grid represents a fibroglandular cluster, and its size represents the median size of the considered tumours, which is consistent with the 2D heterogenous models in [23–26] in which the lesion size is similar to the size of each fibroglandular cluster.

The different tissues within the breast models are modelled using Debye parameters for malignant tissue and for homogeneous lossy adipose tissue, as established by Lazebnik et al. [28, 29]. For the fibroglandular clusters which account for breast heterogeneity,
Figure 1. Cross-section of the 3D FDTD space lattice partitioned into Total Field (TF), Scattered Field (SF) and UPML regions, for a heterogeneous breast model. In this example, the target, a spiculated tumour located at the centre of the TF, is illuminated by a pulsed plane wave propagating in the $+z$ direction (represented by a dark line) and backscatter is recorded at the first observer location: $(0, 0, -74)$ mm (represented by a filled circle). All four observation points are represented by small circles in the image.

there are three different levels of fibroconnective/glandular tissue, as established in [27]. The dielectric properties are shown in Table 1.

Two different configurations of fibroglandular tissue, with varying levels of complexity, are considered in this paper. The first scenario, represented by Models I and II, have a cluster of fibroglandular tissue at a fixed location within the TF region of the breast model. For the second scenario, the best feature extraction and classifier methods are selected from the results from the first scenario and are applied to Models III and IV. Models III and IV include one and two clusters of fibroglandular tissue, respectively, at randomly and independently located within the TF region of the breast model. Further detail on the two scenarios is detailed in the following sections.
2.1. Modelling with Fixed Fibroglandular Tissue Location

For this part of the results, there are two different Models (I and II) in which a cluster of fibroglandular tissue is positioned at a fixed location within the breast, as in studies by Chen et al. [23–26].

For Model I, the portion of heterogeneous breast tissue is located within the cubic TF region in one of its vertices. For Model II, the same block of heterogeneous breast tissue is also located within the TF region, at a distance of $\sqrt[3]{5} \text{mm}$ from one vertex of the cubic TF region, i.e. the block is moved 5 mm in each of the X, Y and Z axes towards the centre of the TF region. In Figure 2, a representation of a sample of benign tumours in Model I and Model II is shown.

For Models I and II, the feature extraction method (PCA) and the classification approach (SVM) are applied. The Radial Basis

Table 1. Debye parameters for the FDTD model. Parameters are established in: (1) Lazebnik et al. [28,29] and (2) UWCEM website [27].

<table>
<thead>
<tr>
<th>Tissue</th>
<th>$\epsilon_\infty$</th>
<th>$\Delta \epsilon$</th>
<th>$\sigma_s \text{ (}\text{sm}^{-1}\text{)}$</th>
<th>$\tau \text{ (ps)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lossy adipose tissue (1)</td>
<td>3.140</td>
<td>1.708</td>
<td>0.036</td>
<td>14.65</td>
</tr>
<tr>
<td>Fibroconnective/glandular — low (2)</td>
<td>9.941</td>
<td>26.60</td>
<td>0.462</td>
<td>10.90</td>
</tr>
<tr>
<td>Fibroconnective/glandular — median (2)</td>
<td>7.821</td>
<td>41.48</td>
<td>0.713</td>
<td>10.66</td>
</tr>
<tr>
<td>Fibroconnective/glandular — high (2)</td>
<td>6.151</td>
<td>48.2</td>
<td>0.809</td>
<td>10.26</td>
</tr>
<tr>
<td>Malignant tissue (1)</td>
<td>6.749</td>
<td>50.09</td>
<td>0.794</td>
<td>10.50</td>
</tr>
</tbody>
</table>

Figure 2. Samples of different Gaussian Random Spheres representing benign tumours. The (a) smooth and (b) macrolobulated tumour models are represented in Model I and Model II, respectively.
Figure 3. Classification architectures in which a 2-stage fine size classification is applied before shape classification: FSCS and FSFS.

Function parameters, $C$ and $\gamma$, for SVM are the same as those used in previous studies by the authors [18, 20], so that tumour classification in heterogeneous breast models can be tested. The following classification architectures were utilised, as these were the ones found to perform the best in the previous papers [17, 18, 20, 21]: Fine-Size Coarse-Shape (FSCS) and Fine-Size Fine-Shape (FSFS). It must be noted that each classification architecture consists of two “partial” sub-classifiers, one for size and one for shape. In Figure 3, a block diagram of the FSCS and FSFS classification architectures are shown.

2.2. Modelling with Varying Fibroglandular Tissue Location

For Models III and IV, the tumour database includes 160 models: i.e., 10 tumour models for each of the four sizes and four shapes.

For Model III the cluster of heterogeneous breast tissue is randomly located in one of ten locations spread within the TF region. For Model IV two independent clusters of heterogeneous breast tissue are randomly located in one of ten random pairs of locations. In Figure 4, a representation of a sample of malignant tumours in Model III and Model IV is shown.

For the second part of the results, PCA and SVM were also used with the same two multi-stage classification architectures (FSCS and FSFS).
Table 2. Classification performance for Models I and II.

<table>
<thead>
<tr>
<th>Architectures of classifiers</th>
<th>Model</th>
<th>Partial size classifier (%)</th>
<th>Partial shape classifier (%)</th>
<th>Size-then-shape classifier (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fine-Size-Coarse-Shape</td>
<td>I</td>
<td>85.83</td>
<td>92.71</td>
<td>79.57</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>85.62</td>
<td>91.25</td>
<td>78.13</td>
</tr>
<tr>
<td>Fine-Size-Fine-Shape</td>
<td>I</td>
<td>85.83</td>
<td>73.96</td>
<td>63.48</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>85.62</td>
<td>75.42</td>
<td>64.57</td>
</tr>
</tbody>
</table>

Figure 4. Samples of different Gaussian Random Spheres representing malignant tumours. The (a) microlobulated and (b) spiculated (10 spicules) tumour models are represented in samples of Model III and Model IV, respectively.

3. RESULTS AND DISCUSSION

3.1. Classification with Fixed Location of Fibroglandular Tissue

The results for Models I and II are shown in Table 2. In this table, partial size and shape classifiers provide classification of tumours in terms of either size or shape only, respectively, while the size-then-shape classifiers attempt to classify a tumour in both size and shape. Finally, the best performances for each stage of the classification architectures are in boldface.

Comparing the results of Table 2 with the previous authors’ homogeneous studies [17, 18, 20], it is observed that the introduction of a portion of fibroglandular breast tissue at a fixed location within the breast model does not significantly degrade the classification performance, suggesting that the algorithm is efficient under these specific conditions.

It is also observed that there is no significant difference between the two fibroglandular clusters location within the breast considered in
Table 3. Classification performance for Models III and IV (one and two randomly-located heterogeneous clusters, respectively).

<table>
<thead>
<tr>
<th>Architectures of classifiers</th>
<th>Number of heterogeneous clusters (Model)</th>
<th>Partial size classifier (%)</th>
<th>Partial shape classifier (%)</th>
<th>Size-then-shape classifier (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fine-Size-Coarse-Shape</td>
<td>one (III)</td>
<td>83.12</td>
<td>90.62</td>
<td>75.33</td>
</tr>
<tr>
<td></td>
<td>two (IV)</td>
<td>80.00</td>
<td>85.00</td>
<td>68.00</td>
</tr>
<tr>
<td>Fine-Size-Fine-Shape</td>
<td>one (III)</td>
<td>83.12</td>
<td>68.12</td>
<td>56.63</td>
</tr>
<tr>
<td></td>
<td>two (IV)</td>
<td>80.00</td>
<td>61.25</td>
<td>49.00</td>
</tr>
</tbody>
</table>

In this study. In the fine shape classifiers, the classification performance difference between Models I and II is 1.46%. For the coarse shape classifiers, the difference of classification performance between the two models is also 1.46%.

3.2. Classification with Varying Fibroglandular Tissue Location

In Models III and IV, the fibroglandular tissue does not have a fixed position within the breast model. The performance results of the classifiers for Models III and IV are shown in Table 3, where the best performances for each stage of the classification architectures are highlighted. As noted earlier, in this section only results for SVM combined with PCA are analysed.

Table 3 shows that a performance decrease is observed when classifying the shape of tumours embedded in breast models with varying locations of fibroglandular tissue, in Model III. It is also observed that there is further shape classification performance decrease when the number of fibroglandular clusters increases to two (as in Model IV). The difference between coarse shape classifiers, where benign and malignant tumours are divided, is slightly lower by 5.62% (90.62% versus 85.00%). This is particularly noticeable in the fine shape classifiers, in which the fine shape of tumours is detected, as the difference can be as high as 6.87% (68.12% versus 61.25% for FSFS). It is also worth noting that the overall size-then-shape classifier failed to classify tumours in both size and shape in Model IV as the performance was only 49%.

For the results presented in both Sections 3.1 and 3.2, the number of misclassified tumours in terms of size is recorded at each step of the two classification architectures. These results illustrate the potential
of the proposed approach to classify the smallest lesions possible, as this compromises the effectiveness of treatment. Table 4 shows the percentage of correctly classified tumours for each of the four sizes.

**Table 4.** Percentage of tumours correctly classified in terms of radius size in Models I, II, III and IV.

<table>
<thead>
<tr>
<th>Models</th>
<th>Percentage of Tumours Correctly Classified in Terms of Radius Size (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coarse Size Classifier</td>
</tr>
<tr>
<td></td>
<td>2.5 and 5 mm</td>
</tr>
<tr>
<td>I</td>
<td>94.58</td>
</tr>
<tr>
<td>II</td>
<td>94.17</td>
</tr>
<tr>
<td>III</td>
<td>96.25</td>
</tr>
<tr>
<td>IV</td>
<td>93.75</td>
</tr>
</tbody>
</table>

In the table above, it can be observed that the coarse size of the tumours is always classified with accuracy over 90%. As far as fine size classification is concerned, the classifier is more accurate with the smallest tumours, 2.5 and 5 mm. In particular, 2.5 mm tumours are classified always with accuracy over 92.50% disregarding the breast model in which tumours are embedded. It must also be noted that while largest tumours are coarsely classified with accuracy of at least 90%, classifying the largest tumours into fine classes of 7.5 and 10 mm result in poorer classification when they are embedded in breast models with randomly located fibroglandular clusters, Models III and IV. In particular 40% of the 7.5 mm tumours are missed by the fine classifier when embedded in Model IV.

**4. CONCLUSION**

The performance of classification algorithms in a dielectrically heterogeneous breast is investigated. A database of tumours with varying size and shape is classified in two different scenarios, with variations regarding breast heterogeneity, using a feature extraction method and a classification algorithm with two multi-stage classification architectures. Four breast models were considered with varying levels of heterogeneity.

A feature extraction method, PCA, was used to extract the most significant features. A SVM classifier with two classification architectures (FSCS and FSFS) were used to classify the tumours.
For the first scenario with fibroglandular tissue fixed in one location, SVM was the most accurate classifier for both partial size and shape classifiers, being the more robust to breast heterogeneity. For Model I, the coarse and fine shape classification achieved a performance as high as 92.71% and 73.96%, respectively. For Model II the coarse and fine shape classification achieved a performance as high as 91.25% and 75.42%, respectively.

For Model III, where the fibroglandular tissue randomly positioned in one of ten different locations within the homogeneous breast model, the coarse and fine shape classification achieved a performance as high as 90.62% and 68.12%, respectively. Finally, for Model IV (two fibroglandular clusters randomly located within the breast), the coarse and fine shape classification achieved a performance as high as 85.00% and 61.25%, respectively.

Overall, the best classification accuracy results are comparable, independent of the breast model in which the tumours are embedded. In fact, the introduction of a fixed structure of fibroglandular tissue does not introduce any significant changes in the classification performance when compared to results for a homogeneous breast model. However, the introduction of cluster(s) of fibroglandular tissue in varying locations within the breast produces a slight decrease in the system performance, more noticeable when two varying fibroglandular clusters are introduced in the breast model at once.

The results presented here are promising for breast tumour classification, within the context of UWB radar imaging, since the feature extraction methods and the classification methods previously analysed have shown to be relatively robust to different tumour sizes, and with the inclusion of clusters of fibroglandular tissue in fixed locations or varying locations within the breast. Future work may include the inclusion of multiple clusters of fibroglandular tissue within the breast model for even more realistic simulations.

**ACKNOWLEDGMENT**

This work is supported by Science Foundation Ireland (SFI) under grant number 07/RFP/ENEF420.

**REFERENCES**

1. Hagness, S. C., A. Taflove, and J. E. Bridges, “Two dimensional FDTD analysis of a pulsed microwave confocal system for breast cancer detection: Fixed-focus and antenna-array sensors,” *IEEE*
12. Byrne, D., M. O’Halloran, E. Jones, and M. Glavin, “Transmitter-


